

日本国特許庁
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されて
いる事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed
with this Office.

出願年月日 2003年 2月19日
Date of Application:

出願番号 特願2003-041833
Application Number:

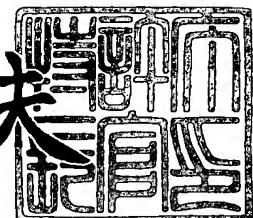
[ST. 10/C] : [JP2003-041833]

出願人 ソニー株式会社
Applicant(s):

2003年12月17日

特許庁長官
Commissioner,
Japan Patent Office

今井康夫



【書類名】 特許願

【整理番号】 0290801003

【提出日】 平成15年 2月19日

【あて先】 特許庁長官 太田 信一郎 殿

【国際特許分類】 G06N 3/00

【発明者】

【住所又は居所】 東京都品川区北品川6丁目7番35号ソニー株式会社内

【氏名】 青山 一美

【発明者】

【住所又は居所】 東京都品川区北品川6丁目7番35号ソニー株式会社内

【氏名】 下村 秀樹

【特許出願人】

【識別番号】 000002185

【氏名又は名称】 ソニー株式会社

【代理人】

【識別番号】 100082740

【弁理士】

【氏名又は名称】 田辺 恵基

【手数料の表示】

【予納台帳番号】 048253

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 9709125

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 学習装置及び学習方法並びにロボット装置

【特許請求の範囲】

【請求項 1】

ユーザとの対話を通して対象とする物体の名前を上記ユーザから取得する対話手段と、

それぞれ上記対象とする物体の異なる所定の特徴を検出すると共に、当該検出結果及び予め記憶している既知の上記物体の対応する上記特徴の学習結果に基づいて、当該対象とする物体を認識する複数の認識手段と、

上記既知の物体の上記名前及び当該既知の物体に対する各上記認識手段の認識結果を関連付けた関連付け情報を記憶する記憶手段と、

上記対話手段が取得した上記対象とする物体の名前、当該対象とする物体に対する各上記認識手段の認識結果、及び上記記憶手段が記憶する上記関連付け情報に基づいて、上記対象とする物体が新規であると判断したときには、必要な上記認識手段に当該対象とする物体の上記対応する特徴を学習させると共に、当該対象とする物体についての上記関連付け情報を上記記憶手段に新たに記憶させる制御手段と

を具えることを特徴とする学習装置。

【請求項 2】

上記制御手段は、

上記既知の物体に対する各上記認識手段の学習達成度を管理し、

上記対話手段が取得した上記対象とする物体の名前、当該対象とする物体に対する各上記認識手段の認識結果、及び上記記憶手段が記憶する上記関連付け情報に基づいて、上記対象とする物体が既知であると判断したときには、当該対象とする物体に対する各上記認識手段の上記学習達成度に基づき決定される所定の上記認識手段に、当該対象とする物体の上記対応する特徴を学習させると共に、学習させた上記認識手段の新たな上記学習達成度に応じて対応する上記関連付け情報を更新する

ことを特徴とする請求項 1 に記載の学習装置。

【請求項 3】

各上記認識手段は、

上記対話手段が上記対象とする物体の上記名前を取得する前から、当該対象とする物体の上記対応する特徴を学習するための学習用データの収集を開始し、当該収集した学習用データを利用して、上記対象とする物体の上記対応する特徴を学習する

ことを特徴とする請求項 1 に記載の学習装置。

【請求項 4】

各上記認識手段は、

所定量の上記学習用データを収集できなかったときにも、当該収集した上記学習用データを利用して上記対象とする物体の上記対応する特徴を学習する
ことを特徴とする請求項 3 に記載の学習装置。

【請求項 5】

各上記認識手段は、

上記対象とする物体の上記対応する特徴の学習を途中で終了した場合に、当該途中までの上記学習結果を保持し、次回の当該対象とする物体の上記対応する特徴の学習を当該途中から開始する

ことを特徴とする請求項 1 に記載の学習装置。

【請求項 6】

上記対象とする物体は、対話相手の上記ユーザであり、

上記対話手段は、

上記ユーザに対する一部又は全部の上記認識手段の学習が不十分であるときに
、上記ユーザとの対話を引き伸ばすための処理を実行する
ことを特徴とする請求項 1 に記載の学習装置。

【請求項 7】

上記対話手段は、

上記ユーザとの対話を引き伸ばす際には、上記対象とする物体に対する上記学習が不十分な上記認識手段が学習し易いような対話を生成するための処理を実行する

ことを特徴とする請求項 6 に記載の学習装置。

【請求項 8】

上記制御手段は、

上記対象とする物体に対する各上記認識手段の上記学習達成度に基づき決定される所定の上記認識手段に、当該対象とする物体の上記対応する特徴を、当該対象とする物体に対する当該認識手段の現在の上記学習達成度に応じた状態から開始するように学習させる

ことを特徴とする請求項 2 に記載の学習装置。

【請求項 9】

上記制御手段は、

各上記認識手段の現在の上記学習達成度を上記記憶手段に記憶させて管理し、上記制御手段又は上記記憶手段は、

上記学習達成度を時間減衰させる

ことを特徴とする請求項 8 に記載の学習装置。

【請求項 10】

ユーザとの対話を通して当該対象とする物体の名前を上記ユーザから取得すると共に、上記対象とする物体の複数の異なる所定の特徴についての検出結果及び予め記憶している既知の上記物体の各上記特徴についての学習結果に基づいて、当該対象とする物体を認識する第 1 のステップと、

取得した上記対象とする物体の名前と、当該対象とする物体の各上記特徴にそれぞれ基づく認識結果と、予め記憶している上記既知の物体の上記名前、当該既知の物体の各上記特徴に対する認識結果を関連付けた関連付け情報に基づいて、上記対象とする物体が新規であると判断したときには、当該対象とする物体の必要な上記特徴を学習すると共に、当該対象とする物体についての上記関連付け情報を新たに記憶する第 2 のステップと

を具えることを特徴とする学習方法。

【請求項 11】

上記第 1 のステップでは、

上記既知の物体の各上記特徴に対する学習達成度を管理し、

上記第2のステップでは、

取得した上記対象とする物体の名前、当該対象とする物体の各上記特徴に対する認識結果、及び記憶している上記関連付け情報に基づいて、上記対象とする物体が既知であると判断したときには、当該対象とする物体の各上記特徴に対する上記学習達成度に基づき決定される所定の上記特徴を学習すると共に、当該特徴に対する新たな上記学習達成度に応じて対応する上記関連付け情報を更新することを特徴とする請求項10に記載の学習方法。

【請求項12】

上記第1のステップでは、

上記対象とする物体の上記名前を取得する前から、当該対象とする物体の各上記特徴を学習するための学習用データの収集を開始し、

上記第2のステップでは、

当該収集した学習用データを利用して、上記対象とする物体の各上記特徴を学習する

ことを特徴とする請求項10に記載の学習方法。

【請求項13】

上記第2のステップでは、

所定量の上記学習用データを収集できなかったときにも、当該収集した上記学習用データを利用して上記対象とする物体の上記対応する特徴を学習する
ことを特徴とする請求項12に記載の学習方法。

【請求項14】

上記第2のステップでは、

上記対象とする物体の上記対応する特徴の学習を途中で終了した場合に、当該途中までの上記学習結果を保持し、

次の当該対象とする物体の上記対応する特徴の学習を当該途中から開始する
ことを特徴とする請求項10に記載の学習方法。

【請求項15】

上記対象とする物体は、対話相手の上記ユーザであり、

上記第2のステップでは、

上記ユーザの一部又は全部の上記特徴に対する学習が不十分であるときに、上記ユーザとの対話を引き伸ばす

ことを特徴とする請求項10に記載の学習方法。

【請求項16】

上記ユーザとの対話を引き伸ばす際には、上記対象とする物体の上記学習が不十分な上記特徴を学習し易いような対話を生成する
ことを特徴とする請求項15に記載の学習方法。

【請求項17】

上記第2のステップでは、

上記対象とする物体の各上記特徴に対する上記学習達成度に基づき決定される所定の上記特徴を、当該特徴に対する現在の上記学習達成度に応じた状態から開始するように学習する

ことを特徴とする請求項11に記載の学習方法。

【請求項18】

上記学習達成度を時間減衰させる

ことを特徴とする請求項11に記載の学習方法。

【請求項19】

ユーザとの対話を通して対象とする物体の名前を上記ユーザから取得する対話手段と、

それぞれ上記対象とする物体の異なる所定の特徴を検出すると共に、当該検出結果及び予め記憶している既知の上記物体の対応する上記特徴の学習結果に基づいて、当該対象とする物体を認識する複数の認識手段と、

上記既知の物体の上記名前及び当該既知の物体に対する各上記認識手段の認識結果を関連付けた関連付け情報を記憶する記憶手段と、

上記対話手段が取得した上記対象とする物体の名前、当該対象とする物体に対する各上記認識手段の認識結果、及び上記記憶手段が記憶する上記関連付け情報に基づいて、上記対象とする物体が新規であると判断したときには、必要な上記認識手段に当該対象とする物体の上記対応する特徴を学習させると共に、当該対象とする物体についての上記関連付け情報を上記記憶手段に新たに記憶させる制

御手段と

を具えることを特徴とするロボット装置。

【請求項 20】

上記制御手段は、

上記既知の物体に対する各上記認識手段の学習達成度を管理し、

上記対話手段が取得した上記対象とする物体の名前、当該対象とする物体に対する各上記認識手段の認識結果、及び上記記憶手段が記憶する上記関連付け情報に基づいて、上記対象とする物体が既知であると判断したときには、当該対象とする物体に対する各上記認識手段の上記学習達成度に基づき決定される所定の上記認識手段に、当該対象とする物体の上記対応する特徴を学習させると共に、学習させた上記認識手段の新たな上記学習達成度に応じて対応する上記関連付け情報を更新する

ことを特徴とする請求項 19 に記載のロボット装置。

【請求項 21】

各上記認識手段は、

上記対話手段が上記対象とする物体の上記名前を取得する前から、当該対象とする物体の上記対応する特徴を学習するための学習用データの収集を開始し、当該収集した学習用データを利用して、上記対象とする物体の上記対応する特徴を学習する

ことを特徴とする請求項 19 に記載のロボット装置。

【請求項 22】

各上記認識手段は、

所定量の上記学習用データを収集できなかったときにも、当該収集した上記学習用データを利用して上記対象とする物体の上記対応する特徴を学習する

ことを特徴とする請求項 21 に記載のロボット装置。

【請求項 23】

各上記認識手段は、

上記対象とする物体の上記対応する特徴の学習を途中で終了した場合に、当該途中までの上記学習結果を保持し、次回の当該対象とする物体の上記対応する特

徴の学習を当該途中から開始する

ことを特徴とする請求項19に記載のロボット装置。

【請求項24】

上記対象とする物体は、対話相手の上記ユーザであり、

上記対話手段は、

上記ユーザに対する一部又は全部の上記認識手段の学習が不十分であるときに

、上記ユーザとの対話を引き伸ばすための処理を実行する

ことを特徴とする請求項19に記載のロボット装置。

【請求項25】

上記対話手段は、

上記ユーザとの対話を引き伸ばす際には、上記対象とする物体に対する上記学
習が不十分な上記認識手段が学習し易いような対話を生成するための処理を実行
する

ことを特徴とする請求項24に記載のロボット装置。

【請求項26】

上記制御手段は、

上記対象とする物体に対する各上記認識手段の上記学習達成度に基づき決定さ
れる所定の上記認識手段に、当該対象とする物体の上記対応する特徴を、当該対
象とする物体に対する当該認識手段の現在の上記学習達成度に応じた状態から開
始するように学習させる

ことを特徴とする請求項20に記載のロボット装置。

【請求項27】

上記制御手段は、

各上記認識手段の現在の上記学習達成度を上記記憶手段に記憶させて管理し、
上記制御手段又は上記記憶手段は、

上記学習達成度を時間減衰させる

ことを特徴とする請求項28に記載のロボット装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は学習装置及び学習方法並びにロボット装置に関し、例えばエンターテインメントロボットに適用して好適なものである。

【0002】**【従来の技術】**

従来、例えばセキュリティシステムなどにおける指紋認識器や声紋認識器等の認識器が新しいカテゴリを学習しようとする場合、学習していることがユーザに明示的であるような「学習モード」といったものが存在する。

【0003】

かかる「学習モード」は、上述のようなセキュリティのために指紋や声紋などを覚えるというような場合には、センシング情報がセキュリティに用いられるという目的がはっきりしているため、今現在自己の情報が登録されているとユーザに知らせるためにも明示的であるほうが好ましい。

【0004】**【発明が解決しようとする課題】**

ところが、インタラクションの最中にユーザが誰であるかを認識するためにセンシング情報を利用するようなエンターテインメントロボットでは、自然なインタラクションの中でユーザを認識できるようにすることが重要となる。

【0005】

このためかかるエンターテインメントロボットにおいて、例えばユーザの顔を学習する際に『顔を覚えるのでじっとしていてください』などとロボットに発話させることにより当該ロボットが顔の学習をしていることをユーザに明示的に示すことは、本来の目的であるユーザとの自然なインタラクションを阻害する問題がある。

【0006】

一方、ユーザの名前を学習するエンターテインメントロボットにおいて、ユーザとの自然なインタラクションを行わせるためには、ユーザから教えてもらった当該ユーザの名前に対してその特徴（その名前と結びつくセンシング情報）をなるべく1回で覚えられるようにする工夫が必要となる。

【0007】

ところが、従来提案されているエンターテインメントロボットでは、一定時間内に十分なデータが得られないときには学習が失敗、というように学習の成否判断が固定的になっているため、動的な環境では学習の失敗が多発し、センシング情報と名前とがなかなか結びつかずにロボットが何度も名前を尋ねてしまうという、ユーザにとって煩わしいインタラクションが発生する問題があった。

【0008】

さらに、従来提案されているエンターテインメントロボットでは、学習データを十分に取得できないまま学習を終了しなければならない場合、そのときの学習は失敗として、その学習により得られたデータは全て廃棄されていた。このためその途中までの学習結果を活かせず、効率の良い学習を行い難い問題があった。

【0009】

従って、エンターテインメントロボットにおいて、自然なインタラクションの中でユーザを認識できるようにすることができ、また学習の失敗という状況をなるべく少なくすることができれば、エンターテインメント性をより一層と向上させ得るものと考えられる。

【0010】

本発明は以上の点を考慮してなされたもので、エンターテインメント性を向上させ得る学習装置及び学習方法並びにロボット装置を提案しようとするものである。

【0011】**【課題を解決するための手段】**

かかる課題を解決するため本発明においては、学習装置において、ユーザとの対話を通じて対象とする物体の名前をユーザから取得する対話手段と、それぞれ対象とする物体の異なる所定の特徴を検出すると共に、当該検出結果及び予め記憶している既知の物体の対応する特徴の学習結果に基づいて、当該対象とする物体を認識する複数の認識手段と、既知の物体の名前及び当該既知の物体に対する各認識手段の認識結果を関連付けた関連付け情報を記憶する記憶手段と、対話手段が取得した対象とする物体の名前、当該対象とする物体に対する各認識手段の

認識結果、及び記憶手段が記憶する関連付け情報に基づいて、対象とする物体が新規であると判断したときには、必要な認識手段に当該対象とする物体の対応する特徴を学習させると共に、当該対象とする物体についての関連付け情報を記憶手段に新たに記憶させる制御手段とを設けるようにした。この結果、この学習装置は、ユーザにそれと認識されることなく、対象とする物体の名前を通常のユーザとの対話を通して学習することができる。

【0012】

また本発明においては、学習方法において、ユーザとの対話を通して当該対象とする物体の名前をユーザから取得すると共に、対象とする物体の複数の異なる所定の特徴についての検出結果及び予め記憶している既知の物体の各特徴についての学習結果に基づいて、当該対象とする物体を認識する第1のステップと、取得した対象とする物体の名前と、当該対象とする物体の各特徴にそれぞれ基づく認識結果と、予め記憶している既知の物体の名前、当該既知の物体の各特徴に対する認識結果を関連付けた関連付け情報を基づいて、対象とする物体が新規であると判断したときには、当該対象とする物体の必要な特徴を学習すると共に、当該対象とする物体についての関連付け情報を新たに記憶する第2のステップとを設けるようにした。この結果、この学習方法によれば、ユーザにそれと認識されることなく、対象とする物体の名前を通常のユーザとの対話を通して学習することができる。

【0013】

さらに本発明においては、ロボット装置において、ユーザとの対話を通して対象とする物体の名前をユーザから取得する対話手段と、それぞれ対象とする物体の異なる所定の特徴を検出すると共に、当該検出結果及び予め記憶している既知の物体の対応する特徴の学習結果に基づいて、当該対象とする物体を認識する複数の認識手段と、既知の物体の名前及び当該既知の物体に対する各認識手段の認識結果を関連付けた関連付け情報を記憶する記憶手段と、対話手段が取得した対象とする物体の名前、当該対象とする物体に対する各認識手段の認識結果、及び記憶手段が記憶する関連付け情報を基づいて、対象とする物体が新規であると判断したときには、必要な認識手段に当該対象とする物体の対応する特徴を学習さ

せると共に、当該対象とする物体についての関連付け情報を記憶手段に新たに記憶させる制御手段とを設けるようにした。この結果、このロボット装置は、ユーザーにそれと認識されることなく、対象とする物体の名前を学習することができる。

【0014】

【発明の実施の形態】

以下図面について、本発明の一実施の形態を詳述する。

【0015】

(1) 本実施の形態によるロボットの構成

図1及び図2において、1は全体として本実施の形態による2足歩行型のロボットを示し、胴体部ユニット2の上部に頭部ユニット3が配設されると共に、当該胴体部ユニット2の上部左右にそれぞれ同じ構成の腕部ユニット4A、4Bがそれぞれ配設され、かつ胴体部ユニット2の下部左右にそれぞれ同じ構成の脚部ユニット5A、5Bがそれぞれ所定位置に取り付けられることにより構成されている。

【0016】

胴体部ユニット2においては、体幹上部を形成するフレーム10及び体幹下部を形成する腰ベース11が腰関節機構12を介して連結することにより構成されており、体幹下部の腰ベース11に固定された腰関節機構12の各アクチュエータA₁、A₂をそれぞれ駆動することによって、体幹上部を図3に示す直交するロール軸13及びピッチ軸14の回りにそれぞれ独立に回転させ得るようになされている。

【0017】

また頭部ユニット3は、フレーム10の上端に固定された肩ベース15の上面中央部に首関節機構16を介して取り付けられており、当該首関節機構16の各アクチュエータA₃、A₄をそれぞれ駆動することによって、図3に示す直交するピッチ軸17及びヨー軸18の回りにそれぞれ独立に回転させ得るようになされている。

【0018】

さらに各腕部ユニット4A、4Bは、それぞれ肩関節機構19を介して肩ベース15の左右に取り付けられており、対応する肩関節機構19の各アクチュエータA5、A6をそれぞれ駆動することによって図3に示す直交するピッチ軸20及びロール軸21の回りにそれぞれ独立に回転させ得るようになされている。

【0019】

この場合、各腕部ユニット4A、4Bは、それぞれ上腕部を形成するアクチュエータA7の出力軸に肘関節機構22を介して前腕部を形成するアクチュエータA8が連結され、当該前腕部の先端に手部23が取り付けられることにより構成されている。

【0020】

そして各腕部ユニット4A、4Bでは、アクチュエータA7を駆動することによって前腕部を図3に示すヨー軸24の回りに回転させ、アクチュエータA8を駆動することによって前腕部を図3に示すピッチ軸25の回りにそれぞれ回転させることができるようになされている。

【0021】

これに対して各脚部ユニット5A、5Bにおいては、それぞれ股関節機構26を介して体幹下部の腰ベース11にそれぞれ取り付けられており、それぞれ対応する股関節機構26の各アクチュエータをA9～A11それぞれ駆動することによって、図3に示す互いに直交するヨー軸27、ロール軸28及びピッチ軸29の回りにそれぞれ独立に回転させ得るようになされている。

【0022】

この場合各脚部ユニット5A、5Bは、それぞれ大腿部を形成するフレーム30の下端に膝関節機構31を介して下腿部を形成するフレーム32が連結されると共に、当該フレーム32の下端に足首関節機構33を介して足部34が連結されることにより構成されている。

【0023】

これにより各脚部ユニット5A、5Bにおいては、膝関節機構31を形成するアクチュエータA12を駆動することによって、下腿部を図3に示すピッチ軸35の回りに回転させることができ、また足首関節機構33のアクチュエータA1

3、A₁4をそれぞれ駆動することによって、足部34を図3に示す直交するピッチ軸36及びロール軸37の回りにそれぞれ独立に回転させ得るようになされている。

【0024】

一方、胴体部ユニット2の体幹下部を形成する腰ベース11の背面側には、図4に示すように、当該ロボット1全体の動作制御を司るメイン制御部40と、電源回路及び通信回路などの周辺回路41と、バッテリ45（図5）となどがボックスに収納されてなる制御ユニット42が配設されている。

【0025】

そしてこの制御ユニット42は、各構成ユニット（胴体部ユニット2、頭部ユニット3、各腕部ユニット4A、4B及び各脚部ユニット5A、5B）内にそれぞれ配設された各サブ制御部43A～43Dと接続されており、これらサブ制御部43A～43Dに対して必要な電源電圧を供給したり、これらサブ制御部43A～43Dと通信を行うことができるようになされている。

【0026】

また各サブ制御部43A～43Dは、それぞれ対応する構成ユニット内の各アクチュエータA₁～A₁₄と接続されており、当該構成ユニット内の各アクチュエータA₁～A₁₄をメイン制御部40から与えられる各種制御コマンドに基づいて指定された状態に駆動し得るようになされている。

【0027】

さらに頭部ユニット3には、図5に示すように、このロボット1の「目」として機能するCCD（Charge Coupled Device）カメラ50及び「耳」として機能するマイクロホン51などの各種外部センサと、「口」として機能するスピーカ52となどがそれぞれ所定位置に配設され、手部23や足部34の底面部等には、外部センサとしてのタッチセンサ53が配設されている。さらに制御ユニット42内には、バッテリセンサ54及び加速度センサ55などからなる内部センサが配設されている。

【0028】

そしてCCDカメラ50は、周囲の状況を撮像し、得られた画像信号S1Aを

メイン制御部に送出する一方、マイクロホン51は、各種外部音を集音し、かくして得られた音声信号S1Bをメイン制御部40に送出するようになされている。さらにタッチセンサ53は、ユーザからの物理的な働きかけや、外部との物理的な接触を検出し、検出結果を圧力検出信号S1Cとしてメイン制御部40に送出する。

【0029】

またバッテリセンサ54は、バッテリ45のエネルギー残量を所定周期で検出し、検出結果をバッテリ残量検出信号S2Aとしてメイン制御部40に送出する一方、加速度センサ56は、3軸方向（x軸、y軸及びz軸）の加速度を所定周期で検出し、検出結果を加速度検出信号S2Bとしてメイン制御部40に送出する。

【0030】

メイン制御部部40は、CCDカメラ50、マイクロホン51及びタッチセンサ53等からそれぞれ供給される外部センサ出力としての画像信号S1A、音声信号S1B及び圧力検出信号S1C等と、バッテリセンサ54及び加速度センサ55等からそれぞれ供給される内部センサ出力としてのバッテリ残量検出信号S2A及び加速度検出信号S2B等とに基づいて、ロボット1の周囲及び内部の状況や、外部物体との接触などを判断する。

【0031】

そしてメイン制御部40は、この判断結果と、予め内部メモリ40Aに格納されている制御プログラムと、そのとき装填されている外部メモリ56に格納されている各種制御パラメータとにに基づいて続く行動を決定し、決定結果に基づく制御コマンドを対応するサブ制御部43A～43Dに送出する。この結果、この制御コマンドに基づき、そのサブ制御部43A～43Dの制御のもとに、対応するアクチュエータA1～A14が駆動され、かくして頭部ユニット3を上下左右に揺動させたり、腕部ユニット4A、4Bを上にあげたり、歩行するなどの行動がロボット1により発現されることとなる。

【0032】

またメイン制御部40は、かかる音声信号S1Bに基づく音声認識処理により

ユーザの発話内容を認識し、当該認識に応じた音声信号S3をスピーカ52に与えることにより、ユーザと対話するための合成音声を外部に出力させる。

【0033】

このようにしてこのロボット1においては、周囲及び内部の状況等に基づいて自律的に行動することができ、またユーザと対話することもができるようになされている。

【0034】

(2) 名前学習機能に関するメイン制御部40の処理

(2-1) 名前学習機能に関するメイン制御部40の構成

【0035】

次にこのロボット1に搭載された名前学習機能について説明する。

【0036】

このロボット1には、ユーザとの対話を通してそのユーザの名前を取得し、当該取得した名前を、そのときマイクロホン51やCCDカメラ50の出力に基づいて検出したそのユーザの声の音響的特徴及び顔の形態的特徴の各データと関連付けて記憶すると共に、これら記憶した各データに基づいて新規なユーザの登場を認識し、その新規なユーザの名前や声の音響的特徴及び顔の形態的特徴を上述と同様に取得し記憶するようにして、ユーザの名前を取得していく名前学習機能が搭載されている。なお以下においては、そのユーザの名前と声の音響的特徴及び顔の形態的特徴を対応付けて記憶し終えたユーザを『既知のユーザ』と呼び、記憶し終っていないユーザを『新規なユーザ』と呼ぶものとする。

【0037】

そしてこの名前学習機能は、メイン制御部40における各種処理により実現されている。

【0038】

ここで、かかる名前学習機能に関するメイン制御部40の処理内容を機能的に分類すると、図6に示すように、ユーザが発声した言葉を認識する音声認識部60と、ユーザの声の音響的特徴を検出すると共に当該検出した音響的特徴に基づいてそのユーザを識別して認識する話者認識部61と、ユーザの顔の形態的特徴

を検出すると共に当該検出した形態的特徴に基づいてそのユーザを識別して認識する顔認識部62と、ユーザとの対話制御を含むユーザの名前を学習するための各種制御を司る対話制御部と、既知のユーザの名前、声の音響的特徴及び顔の形態的特徴の関連付けを管理する連想記憶部と、対話制御部63の制御のもとに各種対話用の音声信号S3を生成してスピーカ54（図5）に送出する音声合成部64とに分けることができる。

【0039】

この場合、音声認識部60は、マイクロホン51（図5）からの音声信号S1Bに基づき所定の音声認識処理を実行することにより当該音声信号S1Bに含まれる言葉を単語単位で認識する機能を有し、認識したこれら単語を文字列データD1として対話制御部63に送出する。

【0040】

話者認識部61は、マイクロホン51からの音声信号S1Bに基づき得られる音声データを例えば内部メモリ40A（図5）に記憶保持する機能と、当該記憶保持した音声データ又はマイクロホン51からリアルタイムで与えられる音声信号S1Bに基づき得られる音声データを用いて、ユーザの声の音響的特徴を例えば“Segregation of Speakers for Recognition and Speaker Identification (CH2977-7/91/0000~0873 S1.00 1991 IEEE)”に記載された方法等を利用した所定の信号処理により検出する機能を有する。

【0041】

そして話者認識部61は、この検出した音響的特徴のデータをそのとき記憶している全ての既知のユーザの音響的特徴のデータと順次比較し、そのとき検出した音響的特徴がいずれか既知のユーザの音響的特徴と一致した場合には当該既知のユーザの音響的特徴と対応付けられた当該音響的特徴に固有の識別子（以下、これをSIDと呼ぶ）を対話制御部63に通知する一方、検出した音響的特徴がいずれか既知のユーザの音響的特徴とも一致しなかった場合には、認識不能を意味するSID（=-1）を対話制御部63に通知する。

【0042】

また話者認識部61は、対話制御部63が新規なユーザであると判断したとき

に当該対話制御部63から与えられる新規学習の開始命令及び学習終了命令に応じて、記憶保持した又はリアルタイムで得られる音声データを用いて、その間そのユーザの声の音響的特徴を検出し、当該検出した音響的特徴のデータを新たな固有のSIDと対応付けて記憶すると共に、このSIDを対話制御部63に通知する。

【0043】

さらに話者認識部61は、その後そのユーザに対する追加学習の開始命令及び終了命令が対話制御部63から与えられたときには、記憶保持した又はリアルタイムで得られる音声データを用いて、そのユーザの声の音響的特徴のデータを追加的に収集する追加学習を実行する。

【0044】

さらに話者認識部61は、対話制御部63からユーザを指定してそのユーザの学習達成度について問い合わせがあったときに、これに回答する回答機能を有する。ここで学習達成度とは、そのユーザを認識するために用いるデータ（ここでは音響的特徴のデータ）の収集の程度を意味し、話者認識部61における学習達成度は、そのユーザの声の音響的特徴のデータを収集するために用いた音声の時間長をパラメータとした関数で与えられる値をもとに決定される。

【0045】

そしてこの実施の形態においては、かかる学習達成度として、認識に実用上十分に使用できるレベルである「A（十分に学習できた）」レベルと、認識には使用可能であるが追加学習をした方がいい程度のレベルである「B（ちょっと不安）」レベルと、認識には使用するには不十分であるため認識に使用せず、次の機会に追加学習をすべきレベルである「C（不十分）」レベルとが数値として設定されている。

【0046】

かくして話者認識部61においては、対話制御部63からユーザを指定してそのユーザの学習達成度について問い合わせがあったときには、そのユーザの声の音響的特徴のデータを収集するために用いた音声の時間長をパラメータとした関数で与えられる値から、そのユーザ学習達成度が「A」～「C」のいずれに該当

するかを判定し、判定結果を対話制御部 63 に通知する。

【0047】

顔認識部 62においては、CCDカメラ 50（図5）からの画像信号に基づき得られる学習用の画像データを例えれば内部メモリ 40A（図5）に記憶保持する機能と、当該記憶保持した画像データ又はCCDカメラ 50からリアルタイムで与えられる当該画像信号 S1Aに基づき得られる画像データを用いて、当該画像データに基づく画像内に含まれるユーザの顔の形態的特徴を所定の信号処理により検出する機能とを有する。

【0048】

そして顔認識部 62は、この検出した形態的特徴のデータをそのとき記憶している全ての既知のユーザの顔の形態的特徴のデータと順次比較し、そのとき検出した形態的特徴がいずれか既知のユーザの顔の形態的特徴と一致した場合には当該既知のユーザの形態的特徴と対応付けられた当該形態的特徴に固有の識別子（以下、これをFIDと呼ぶ）を対話制御部 63に通知する一方、検出した形態的特徴がいずれか既知のユーザの顔の形態的特徴とも一致しなかった場合には、認識不能を意味する FID (= -1) を対話制御部 63に通知する。

【0049】

また顔認識部 62は、対話制御部 63が新規なユーザであると判断したときに当該対話制御部 63から与えられる新規学習の開始命令及び終了命令に基づいて、記憶保持した又はリアルタイムで得られる画像データを用いて、その間ユーザの顔の形態的特徴を検出し、当該検出した形態的特徴のデータを新たな固有のFIDと対応付けて記憶すると共に、このFIDを対話制御部 63に通知する。

【0050】

さらに顔認識部 62は、その後そのユーザに対する追加学習の開始命令及び終了命令が対話制御部 63から与えられたときには、記憶保持した又はリアルタイムで得られる画像データを用いて、そのユーザの顔の形態的特徴のデータを追加的に収集する追加学習を実行する。

【0051】

さらに顔認識部 62は、話者認識部 61と同様に、対話制御部 63からユーザ

を指定してそのユーザの学習達成度について問い合わせがあったときに、これに回答する回答機能を有する。そしてこの実施の形態の場合、顔認識部62における学習達成度は、ユーザの顔の形態的特徴のデータを収集するために用いた画像信号S1Aに基づく顔画像の枚数をパラメータとした関数で与えられる値をもとに決定される。

【0052】

かくして顔認識部62は、対話制御部63からユーザを指定してそのユーザの学習達成度について問い合わせがあったときには、かかる値から学習達成度が「A」～「C」のいずれに該当するかを判定し、判定結果を学習達成度として対話制御部63に通知する。

【0053】

音声合成部64においては、対話制御部63から与えられる文字列データD2を音声信号S3に変換する機能を有し、かくして得られた音声信号S3をスピーカ54（図5）に送出することにより、この音声信号S3に基づく音声をスピーカ54から出力させる。

【0054】

連想記憶部65においては、例えば内部メモリ40A（図5）とソフトウェアとからなるオブジェクトであり、対話制御部63の制御のもとに、図7に示すように、既知のユーザの名前と、話者認識部61が記憶しているそのユーザの声の音響的特徴のデータに対応付けられたS1Dと、顔認識部62が記憶しているそのユーザの顔の形態的特徴のデータに対応付けられたF1Dとを記憶する。

【0055】

この際連想記憶部65は、同じユーザと対応する名前、S1D及びF1Dを関連付けて記憶するようになされ、これにより既知のユーザについて、1つの情報（名前、S1D及びF1D）から他の情報を検索し得るようになされている。

【0056】

また連想記憶部65は、対話制御部63の制御のもとに、かかる各既知のユーザのS1Dとそれぞれ対応付けて、そのユーザに対する話者認識部61の学習達成度を記憶すると共に、これと同様にして、各既知のユーザのF1Dとそれぞれ

対応付けて、そのユーザに対する顔認識部62の学習達成度を記憶する。

【0057】

なおこの実施の形態の場合、連想記憶部65は、かかる記憶した話者認識部61及び顔認識部62の各ユーザに対する学習達成度を、その登録後又は最後の更新後一定時間（例えば数日）が経過するごとにレベルを下げる（例えば「A」であったものを「B」に下げ、「B」であったものを「C」に下げる）（以下、これを学習達成度を時間減衰させるという）ように更新するようになされている。

【0058】

これは、ユーザの顔や声が経時的に変化することから、話者認識部61や顔認識部62がユーザを認識するために用いるそのユーザの声の音響的特徴のデータや顔の形態的特徴のデータも一定時間毎に更新することが望ましいからである。

【0059】

対話制御部63は、音声認識部60からの文字列データD1やプログラム等に基づき必要な文字列データD2を音声合成部64に順次与えることにより、ユーザに対して適切な応答や質問を行う対話制御機能を有する。

【0060】

そして対話制御部は、かかる応答や質問を通して取得したそのユーザの名前と、このときのそのユーザに対する話者認識部61及び顔認識部の各認識結果とに基づき、連想記憶部に記憶された各既知のユーザの名前、S I D及びF I D等の情報を参照しながらそのユーザが新規なユーザであるか否かを判断する。

【0061】

そして対話制御部63は、そのユーザが新規なユーザであると判断したときは、話者認識部61及び顔認識部62に対して新規学習の開始命令及び終了命令を与えることにより、これら話者認識部61及び顔認識部62にその新規なユーザの声の音響的特徴や顔の形態的特徴のデータを収集及び記憶（学習）させる。

【0062】

また対話制御部63は、この後話者認識部61及び顔認識部62に対し、所定のタイミングでそのユーザに対する学習達成度を問い合わせ、話者認識部61及び顔認識部62のいずれかから「C」との回答を得た場合には、その話者認識部

6 1 及び又は顔認識部 6 2 に対して追加学習の開始命令及び終了命令を与えることにより、その話者認識部 6 1 及び又は顔認識部 6 2 に追加学習を行わせる一方、その間ユーザとの対話を長引かせるような制御を実行する。

【0063】

これに対して対話制御部 6 3 は、そのユーザが既知のユーザであると判断したときには、連想記憶部 6 5 が記憶しているそのユーザに対する話者認識部 6 1 及び顔認識部 6 2 の学習達成度を確認し、学習達成度が「B」又は「C」の話者認識部 6 1 及び又は顔認識部 6 2 にその学習達成度を通知すると共に、その話者認識部 6 1 及び又は顔認識部 6 2 に対して追加学習の開始命令及び終了命令を与えることにより追加学習を行わせる。

【0064】

そして対話制御部 6 3 は、かかる話者認識部 6 1 及び又は顔認識部 6 2 の追加学習の終了後、その話者認識部 6 1 及び又は顔認識部 6 2 にそのときのそのユーザに対する学習達成度を問い合わせ、これに対する話者認識部 6 1 及び又は顔認識部 6 2 の回答に基づき連想記憶部 6 5 に記憶された対応する学習達成度を更新する。

【0065】

(2-2) 名前学習機能に関する対話制御部 6 3 の具体的処理

次に、名前学習機能に関する対話制御部 6 3 の具体的な処理内容について説明する。

【0066】

対話制御部 6 3 は、内部メモリ 40A に格納された制御プログラムに基づいて、図 8 に示す名前学習処理手順 R T 1 に従って新規な人の名前を順次学習するための各種処理を実行する。

【0067】

すなわち対話制御部 6 3 は、CCD カメラ 50 からの画像信号 S 1 A に基づき顔認識部 6 2 がユーザの顔を認識することにより当該顔認識部 6 2 から FID が与えられると、この名前学習処理手順 R T 1 をステップ S P 0 において開始し、続くステップ S P 1 において、話者認識部 6 1 及び顔認識部 6 2 を制御すること

により、話者認識部61にマイクロホン51（図5）からの音声信号S1Bに基づく音声データの記憶を開始させると共に、顔認識部62にCCDカメラ50からの画像信号S1Aに基づく画像データの記憶を開始させる。

【0068】

続いて対話制御部63は、ステップSP2に進んで、そのユーザの名前を確定する。具体的には、先行して得られたそのユーザのFIDに基づき、連想記憶部65に関連付けて記憶された各既知のユーザの名前、SID及びFIDからそのユーザの名前が検索できるか否かを判断し、検索できた場合には、これに応じた文字列データD2を音声合成部64に送出することにより、例えば「〇〇さんですね。」といったそのユーザの名前がFIDから検索された名前（上述の〇〇に当てはまる名前）と一致するか否かを確かめるための音声を出力させる。

【0069】

そして対話制御部63は、かかる質問に対するそのユーザの「はい、そうです。」といった肯定的な応答の音声認識結果を音声認識部60からの文字列データD1に基づき認識できた場合には、そのユーザの名前をかかる「〇〇さん」に確定する。

【0070】

これに対して対話制御部63は、「いいえ、違います。」といった否定的な応答の音声認識結果を音声認識部60からの文字列データD1に基づき認識できた場合には、これに応じた文字列データD2を音声合成部64に送出することにより、例えば図9に示すように、「名前を教えてください。」といった名前を聞き出すための音声を出力させる。

【0071】

そして対話制御部63は、かかる質問に対するそのユーザの「〇〇です。」といった応答の音声認識結果（すなわち名前）が得られ、その後さらに「〇〇さんっていうんですね。」といった確認に対するユーザの肯定的な応答があったことを音声認識部60からの文字列データD1に基づいて認識すると、そのユーザの名前をかかる「〇〇さん」に確定する。

【0072】

そして対話制御部63は、このようにしてそのユーザの名前を確定すると、この後ステップSP3に進んで、話者認識部61及び顔認識部62を制御することにより、話者認識部61にマイクロホン51からの音声信号S1Aに基づく音声データの記憶処理を終了させると共に、顔認識部62にCCDカメラ50からの画像信号S1Aに基づく画像データの記憶を終了させる。

【0073】

続いて対話制御部63は、ステップSP4に進んで、連想記憶部65が記憶している各既知のユーザの名前と、そのユーザに対応するSID及FIDとに基づいて、ステップSP2において確定したそのユーザの名前と関連付けられたSID及びFIDが存在するか否かを判断する。

【0074】

ここで、このステップSP4で否定結果を得ることは、そのユーザが、話者認識部61がそのユーザの声の音響的特徴のデータを全く収集しておらず、かつ顔認識部62がそのユーザの顔の形態的特徴のデータを全く収集していない新規なユーザであることを意味する。

【0075】

かくしてこのとき対話制御部63は、ステップSP8に進んで、話者認識部61及び顔認識部62に対して、ステップSP1～ステップSP3間において記憶した音声データ又は画像データを利用した新規学習の開始命令を通知する。この結果、これら話者認識部61及び顔認識部62において、かかる音声データ又は画像データを利用して、新たにそのユーザの声の音響的特徴のデータ又はそのユーザの顔の形態的特徴のデータを収集し記憶する新規学習が開始されることとなる。

【0076】

一方、ステップSP4において肯定結果を得ることは、そのユーザが、話者認識部61及び顔認識部62がそれぞれそのユーザの声の音響的特徴のデータ又は顔の形態的特徴のデータを既に収集している既知のユーザであることを意味する。

【0077】

かくしてこのとき対話制御部63は、ステップSP5に進んで、連想記憶部65がそのユーザのSIDと対応付けて記憶しているそのユーザに対する話者認識部61の学習達成度と、連想記憶部65がそのユーザのFIDと対応付けて記憶しているそのユーザに対する顔認識部62の学習達成度とをそれぞれ確認する。

【0078】

ここで、かかる確認結果として、そのユーザに対する話者認識部61及び顔認識部62の学習達成度がいずれも「A」であることが確認された場合には、話者認識部61及び顔認識部62が既にそのユーザの認識に十分な当該ユーザの声の音響的特徴のデータ又は顔の形態的特徴のデータを収集し終えていると判断できる。

【0079】

よって、この場合対話制御部63は、ステップSP6に進んで、話者認識部61及び顔認識部63にステップSP1～ステップSP3間において記憶させた音声データ又は画像データを破棄すべき旨の命令を通知する。

【0080】

また対話制御部63は、この後ステップSP14に進んで、連想記憶部65が記憶しているそのユーザに対する話者認識部61及び顔認識部62の学習達成度を再度「A」に更新した後、ステップSP15に進んでこの名前学習処理手順R T1を終了する。なお、この後ロボット1は、そのユーザに対する学習とは無関係にそのユーザとの対話やダンス等の各種インタラクションを行う。

【0081】

これに対してステップSP5における確認結果として、そのユーザに対する話者認識部61及び顔認識部62のいずれかの一方又は両方の学習達成度が「B」又は「C」であることが確認された場合には、その話者認識部61及び又は顔認識部62が未だそのユーザの認識に十分な当該ユーザの声の音響的特徴のデータ又は顔の形態的特徴のデータを収集し終えていないと判断できる。

【0082】

よって、この場合対話制御部63は、ステップSP7に進んで学習達成度が「B」又は「C」であった話者認識部61及び又は顔認識部62に対してその学習

達成度を通知し、この後ステップSP8に進んで、その学習達成度を通知した（すなわちそのユーザに対する学習達成度が未だ「B」又は「C」である）話者認識部61及び又は顔認識部62に対してステップSP1～ステップSP3間において記憶した音声データ又は画像データを利用した追加学習の開始命令を通知する。

【0083】

この結果、その話者認識部61及び又は顔認識部62において、ステップSP7において通知された学習達成度をスタートポイントとして、そのユーザの声の音響的特徴のデータ又は顔の形態的特徴のデータの収集が未だ十分でない現在の収集状態からさらにこれを収集するように、かかる音声データ又は画像データを利用した追加学習が開始される。

【0084】

さらに対話制御部63は、この後その話者認識部61及び又は顔認識部62からステップSP1～ステップSP3間において記憶した音声データ又は画像データを利用した学習が終了した旨の通知が与えられると、ステップSP9に進んで、その学習を行った話者認識部61及び又は顔認識部62にそのユーザに対する学習達成度を問い合わせ、これに対するその話者認識部61及び又は顔認識部62の回答がともに「A」又は「B」のいずれかであるか否かを判断する。

【0085】

ここで、このステップSP9において肯定結果を得ることは、学習を行った話者認識部61及び又は顔認識部62のいずれもがそのユーザを認識できる程度に当該ユーザの声の音響的特徴のデータ又は顔の形態的特徴のデータを収集し記憶し終えた（十分に学習し終えた）ことを意味する。

【0086】

かくしてこのとき対話制御部63は、ステップSP14に進んで、連想記憶部65が記憶しているそのユーザに対する話者認識部61及び顔認識部62の学習達成度を、学習を行っていないものについてはステップSP5において確認した学習達成度、学習を行ったものについてはステップSP9において得られた学習達成度に更新し、この後ステップSP15に進んでこの名前学習処理手順RT1

を終了する。なお、この後ロボット1は、そのユーザに対する学習とは無関係にそのユーザとの対話やダンス等の各種インタラクションを行う。

【0087】

これに対してステップSP9において否定結果を得ることは、学習を行った話者認識部61及び又は顔認識部62のすくなくとも一方が未だそのユーザを認識できる程度に当該ユーザの声の音響的特徴のデータ又は顔の形態的特徴のデータを収集し記憶していない（十分には学習し終えていない）ことを意味する。

【0088】

かくしてこのとき対話制御部63は、ステップSP10に進んで、その話者認識部61及び又は顔認識部62に対して、マイクロホン51からリアルタイムで与えられる音声信号S1Bに基づき得られる音声データ又はCCDカメラ50からリアルタイムで与えられる画像信号S1Aに基づき得られる画像データを利用した追加学習の開始命令を通知する。

【0089】

この結果、その話者認識部61及び又は顔認識部62において、ステップSP9で対話制御部63に通知した学習達成度をスタートポイントとして、かかる音声データ又は画像データを利用して、そのユーザの声の音響的特徴のデータ又はそのユーザの顔の形態的特徴のデータを追加的に収集し記憶する追加学習が開始される。

【0090】

また対話制御部63は、この後ステップSP11に進んで、そのユーザとの対話を引き伸ばすための処理を実行する。具体的には、対話制御部63は、ロボット1がそのユーザについて学習していることを認識させないように、例えば図10に示すように、「僕とお友達になってくれる?」、「ありがとう!それじゃ、○○さんのこと聞いていい?」、「○○さんの好きなたべものって何?」などのようなロボット1の方からそのユーザに対して積極的に話しかけたり、話題を提供するための文字列データD2を、音声認識部60によるユーザの発話内容の音声認識結果に応じて選択しながら音声合成部64に送出する。

【0091】

また対話制御部63は、この後所定のタイミングでステップSP12に進んで、追加学習の開始命令を通知した話者認識部61及び又は顔認識部62にそのユーザに対する学習達成度を問い合わせ、これに対するその話者認識部61及び又は顔認識部62の回答がともに「A」又は「B」のいずれかであるか否かを判断する。

【0092】

そして対話制御部は、このステップSP12において否定結果を得るとステップSP11に戻り、この後このステップSP12において肯定結果を得られるまでステップSP11～SP12～SP11のループを繰り返す。

【0093】

そして対話制御部63は、やがて話者認識部61及び顔認識部62のいずれもがその後そのユーザを認識できる程度に当該ユーザの声の音響的特徴のデータ又は顔の形態的特徴のデータを収集し記憶し終えることによりステップSP12において肯定結果を得ると、ステップSP13に進んで、追加学習を行っている話者認識部61及び又は顔認識部62に対して追加学習の終了命令を通知する。

【0094】

また対話制御部63は、この後ステップSP14に進んで、連想記憶部65が記憶しているそのユーザに対する話者認識部61及び顔認識部62の学習達成度を、ステップSP10～ステップSP13間の追加学習を行っていないものについてはステップSP5又はステップ9において確認した学習達成度、ステップSP10～ステップSP13間の追加学習を行ったものについてはステップSP12において得られた学習達成度に更新し、この後ステップSP15に進んでこの名前学習処理手順RT1を終了する。なお、この後ロボット1は、そのユーザに対する学習とは無関係にそのユーザとの対話やダンス等の各種インタラクションを行う。

【0095】

このようにしてこのロボット1においては、対話制御部63の制御のもとに、新規なユーザに対する新規学習や、既知のユーザに対する追加学習を行い得、これにより新規なユーザの名前をその声の音響的特徴のデータ及びその顔の形態的

特徴のデータと関連付けて順次学習し得るようになされている。

【0096】

(2-3) 名前学習処理におけるエラー処理

次に、かかる名前学習処理手順S T 1に従った名前学習処理時において、学習中のユーザが立ち去ってしまった等の理由によりそのユーザに対する学習を途中で終了せざるを得ない場合の処理について説明する。

【0097】

ユーザに対する学習を途中で終了せざるを得ない場合としては、

- ①名前学習処理手順S T 1のステップS P 2において、そのユーザの名前を確定する前に学習を終了しなければならない場合
 - ②名前学習処理手順S T 1のステップS P 1～ステップS P 3間において、話者認識部や顔認識部が音声データ又は画像データの記憶保持を開始後、その終了前に学習を終了せざるを得ない場合
 - ③名前学習処理手順S T 1のステップS P 4～ステップS P 7間において、話者認識部6 1や顔認識部6 2が、記憶した音声データや画像データを利用した学習を開始する前に学習を終了せざるを得ない場合
 - ④名前学習処理手順R T 1のステップS P 8～ステップS P 9間において、話者認識部6 1や顔認識部6 2が、記憶した音声データ又は画像データを利用した新規学習又は追加学習中に当該学習を終了せざるを得ない場合
 - ⑤名前学習処理手順R T 1のステップS P 10～ステップS P 13間において、話者認識部6 1や顔認識部6 2が、リアルタイムで得られる音声信号S 1 B又は画像信号S 1 Aを利用した追加学習中に学習を終了せざるを得ない場合
- の5つのパターンが考えられる。なお、以下においては、これらパターンをそれぞれ学習不能パターン①～⑤と呼ぶものとする。

【0098】

これらの場合、そのユーザに対する学習は失敗であるとして、それまでに話者認識部6 1や顔認識部6 2が記憶した学習用の音声データ若しくは画像データを破棄させ、又はそれまでの間に話者認識部6 1や顔認識部6 2が収集したそのユーザの声の音響的特徴のデータや顔の形態的特徴のデータを破棄させることも考

えられるが、これではそれまでに収集した学習用の音声データ又は画像データや、ユーザの声の音響的特徴のデータ又は顔の形態的特徴のデータが無駄となる。

【0099】

そこで、このロボット1においては、上述の学習不能パターン①の場合には、そのときまでに記憶した学習用の音声データや画像データを破棄させる一方、上述の学習不能パターン②又は学習不能パターン③の場合には、そのときまでに記憶した学習用の音声データや画像データを利用して必要な話者認識部61及び又は顔認識部62に学習を行わせ、当該学習により学習達成度が「A」又は「B」となった場合にはかかる学習を有効なものとして取り扱い、未だ学習達成度が「C」の場合にはかかる学習を無効なものとして取り扱う。

【0100】

またこのロボット1においては、上述の学習不能パターン④又は学習不能パターン⑤の場合には、学習を終了せざるを得なくなった時点での話者認識部61や顔認識部62の学習達成度に応じて、当該学習達成度が「A」、「B」である場合にはかかる学習を有効なものとして取り扱い、学習達成度が「C」の場合にはかかる学習を無効なものとして取り扱うようになされている。

【0101】

ここで、このようなエラー処理は、図11に示すエラー処理手順RT2に従って対話制御部63の制御のもとに行われる。

【0102】

実際に、対話制御部63は、上述の名前学習処理手順RT1の実行時に、例えば顔認識部61から認識対象のユーザを検出し得なくなったとのエラー通知が与えられるなど、そのユーザに対する学習を継続し得なくなる予め定められた所定状態が発生すると、名前学習処理手順RT1を終了してこのエラー処理手順RT2をステップSP20において開始し、続くステップSP21において、それまで対象としていたユーザの名前が確定しているか否かを判断する。

【0103】

このステップSP21において否定結果を得ることは、かかるエラーが名前学習処理手順RT1のステップSP2においてユーザの名前を確定する前に学習を

終了しなければならない事態が発生したことを意味し（学習不能パターン①の場合）、このとき対話制御部63は、ステップSP22に進んで、話者認識部61及び顔認識部62に名前学習処理手順RT1の開始後そのときまでに記憶した学習用の音声データ又は画像データを破棄すべき旨の命令を通知し、この後ステップSP34に進んでこのエラー処理手順RT2を終了する。

【0104】

これに対して対話制御部63は、このステップSP21において肯定結果を得ると、ステップSP23に進んで、話者認識部61及び顔認識部62に自己の状態を通知すべき旨の命令を与え、これに対する話者認識部61及び顔認識部62の応答に基づいて、これら話者認識部61及び顔認識部62の現在の状態を確認する。

【0105】

そして対話制御部63は、このステップSP23において、話者認識部61及び顔認識部62が音声データ又は画像データの記憶保持を開始後、その終了前であることを確認すると（学習不能パターン②の場合）、ステップSP24に進んで、話者認識部61及び顔認識部62に学習用の音声データ又は画像データの記憶を終了すべき旨の命令を通知し、この後ステップSP25～ステップSP29を上述の名前学習処理手順RT1のステップSP4～ステップSP8と同様に処理する。

【0106】

そして対話制御部63は、かかるステップSP25～ステップSP29を処理し終えると、ステップSP30に進んで、話者認識部61及び顔認識部62にそのユーザに対する学習達成度を問い合わせることにより、そのユーザに対する話者認識部61及び顔認識部62の学習達成度を取得し、この後ステップSP31に進んで、これら取得したそのユーザに対する話者認識部61及び顔認識部62の学習達成度がともに「C」であるか否かを判断する。

【0107】

ここで、このステップSP31において否定結果を得ることは、かかる話者認識部61及び顔認識部62がともにそのユーザについて、当該ユーザを認識でき

る程度にそのユーザの声の音響的特徴のデータ又は顔の形態的特徴のデータを収集し得たことを意味する。

【0108】

かくしてこのとき対話制御部63は、ステップSP33に進んで、かかるユーザが新規のユーザであった場合には、新たに話者認識部61及び顔認識部62からそれぞれ発行されたS1D及びF1Dと、名前学習処理手順RT1のステップSP2において確定したそのユーザの名前と、このエラー処理手順RT2のステップSP30において取得したそのユーザに対する話者認識部61及び顔認識部62の学習達成度とを上述の様に関連付けて連想記憶部65に記憶させる。

【0109】

また対話制御部63は、かかるユーザが既知のユーザであった場合には、連想記憶部65が記憶しているそのユーザに対する話者認識部61及び顔認識部62の学習達成度を、ステップSP30において取得したそのユーザに対する話者認識部61及び顔認識部62の現在の学習達成度に更新する。そして対話制御部63は、この後ステップSP34に進んでこのエラー処理手順RT2を終了する。

【0110】

これに対して、ステップSP31において否定結果を得ることは、話者認識部61及び顔認識部62のいずれか一方又は両方がそのユーザについて、当該ユーザを認識できる程度にそのユーザの声の音響的特徴のデータ又は顔の形態的特徴のデータを収集し終えていないことを意味する。

【0111】

かくしてこのとき対話制御部63は、ステップSP32に進んで、話者認識部61及び顔認識部62に対して、この後の話者認識処理や認識処理においてそのS1D又はF1Dと対応付けられた声の音響的特徴のデータや顔の形態的特徴のデータを使用しないように指示を与える。この結果、話者認識部61及び顔認識部62において、かかる声の音響的特徴のデータや顔の形態的特徴のデータは、この後の追加学習によりそのユーザに対する学習達成度が「B」又は「A」となるまで話者認識処理又は顔認識処理に使用されないこととなる。

【0112】

そして対話制御部63は、この後ステップS P 3 3に進んで、かかるユーザが新規のユーザであった場合には、新たに話者認識部61及び顔認識部62から発行されたそれぞれS I D及びF I Dと、名前学習処理手順R T 1のステップS P 2において確定したそのユーザの名前と、このエラー処理手順R T 2のステップS P 3 0において取得したそのユーザに対する話者認識部61及び顔認識部62の学習達成度とを上述のように関連付けて連想記憶部65に記憶させる。

【0113】

また対話制御部63は、かかるユーザが既知のユーザであった場合には、連想記憶部65が記憶しているそのユーザに対する話者認識部61及び顔認識部62の学習達成度を、ステップS P 3 0において取得したそのユーザに対する話者認識部61及び顔認識部62の現在の学習達成度に更新する。そして対話制御部63は、この後ステップS P 3 4に進んでこのエラー処理手順R T 2を終了する。

【0114】

従って、このロボット1の場合、そのユーザに対する話者認識部61及び又は顔認識部62の学習達成度が「C」であっても、そのとき得られた当該ユーザの名前と、対応するS I D及びF I Dと、そのユーザに対する話者認識部61及び又は顔認識部62の学習達成度とが関連付けて連想記憶部65に記憶保持されるため（ステップS P 3 1～ステップS P 3 3）、例えば次にその名前をもつユーザを認識したときに例えば図12のように、そのユーザに以前会ったことがあるという発話をを行うことができる。

【0115】

またステップS P 3 2において、対話制御部63から話者認識部61及び又は顔認識部62に対してそのとき収集したユーザの声の音響的特徴のデータや顔の形態的特徴のデータを使用しないように指示が与えられた場合にあっても、名前学習処理手順R T 1のステップS P 7及びステップS P 8について上述したように、次の当該ユーザに対する話者認識部61及び又は顔認識部62の追加学習が当該収集したユーザの声の音響的特徴のデータや顔の形態的特徴のデータの存在を前提とした途中から開始されるため、これら話者認識部61及び顔認識部62が効率良く学習を行うことができる。

【0116】

一方、対話制御部63は、ステップSP23において、話者認識部61及び顔認識部62の現在の状態として、音声データ又は画像データを記憶し終えたが、これを利用した学習を開始する前であることを確認すると（学習不能パターン③の場合）、ステップSP25に進んで、名前学習処理手順RT1のステップSP2において確定したそのユーザの名前に基づいて、当該名前と関連付けられたSID及びFIDを連想記憶部65が記憶しているか否かを判断し、この後ステップSP26～ステップSP34を上述と同様に処理する。

【0117】

また対話制御部63は、ステップSP23において、話者認識部61及び顔認識部62の現在の状態として、記憶した音声データ又は画像データを利用した学習中であることを確認すると、（学習不能パターン④の場合）、ステップSP30に進んで、話者認識部61及び顔認識部62にそのユーザに対する学習達成度を問い合わせることにより、そのユーザに対する話者認識部61及び顔認識部62の学習達成度を取得し、この後ステップSP31～ステップSP34を上述と同様に処理する。

【0118】

さらに対話制御部63は、ステップSP23において、話者認識部61及び顔認識部62の現在の状態として、マイクロホン51からリアルタイムで与えられる音声信号S1Bに基づく音声データ又はCCDカメラ50からリアルタイムで与えられる画像信号S1Aに基づく画像データを利用した追加学習中であることを確認すると（学習不能パターン⑤の場合）、ステップSP35に進んで、その話者認識部61及び又は顔認識部62に追加学習の終了命令を通知する。

【0119】

また対話制御部63は、この後ステップSP30に進んで、話者認識部61及び顔認識部62にそのユーザに対する学習達成度を問い合わせることにより、そのユーザに対する話者認識部61及び顔認識部62の学習達成度を取得し、さらにこの後ステップSP31～ステップSP34を上述と同様に処理する。

【0120】

このようにしてこのロボット1においては、ユーザを学習中にそのユーザの学習を終了せざるを得なくなった場合においても、対話制御部63の制御のもとに、そのときまでに収集した学習用の音声データ又は画像データや、ユーザの声の音響的特徴のデータ又は顔の形態的特徴のデータを利用してそのユーザを行い得るようになされている。

【0121】

(3) 本実施の形態の動作及び効果

以上の構成において、このロボット1では、新規なユーザとの対話を通してそのユーザの名前を取得し、当該名前を、マイクロホン51(図5)やCCDカメラ50(図5)の出力に基づいて検出したそのユーザの声の音響的特徴及び顔の形態的特徴の各データと関連付けて記憶すると共に、これら記憶した各種データに基づいて、名前を取得していないさらに新規なユーザの登場を認識し、その新規なユーザの名前や声の音響的特徴及び顔の形態的特徴を上述と同様にして取得し記憶するようにしてユーザの名前を学習する。

【0122】

従って、このロボット1は、学習していることをそのユーザに認識させることなく、人間が普段行うように、ユーザとの対話を通じて新規のユーザの名前やそのユーザの声の音響的特徴及び顔の形態的特徴を自然に学習することができる。

【0123】

またこのロボット1は、かかる学習を行うに際して話者認識部61及び顔認識部62が、対象とするユーザの声の音響的特徴を学習するための学習用の音声データ又は及びユーザの顔の形態的特徴を学習するための学習用の画像データを当該ユーザの名前が確定する前から予め記憶しておく、この音声データ及び画像データを利用して学習を行うこととしているため、そのユーザに対する学習を途中で終了せざるを得ない事態が発生した場合にもそのユーザに対する学習を行うことができる可能性があり、その分より効率良くユーザの学習を行うことができる。

【0124】

さらにこのロボット1は、ユーザの学習を学習途中で終了せざるを得ない場合

にも、それまでの学習結果であるユーザの声の音響的特徴のデータや顔の形態的特徴のデータを保持し、次のそのユーザに対する学習をその途中の状態から開始するため、効率良く学習を行うことができる。

【0125】

さらにこのロボット1は、予め記憶した対象とするユーザの声の音声データ及び顔の画像データを利用した学習によって十分な学習が行われなかった場合には、そのユーザとの対話を引き伸ばして学習を継続することとしているため、そのユーザに対する学習を1回の対話によって完了させる機会が多く、その分同じユーザに名前を何度も尋ねるというユーザにとって煩わしいインタラクションの発生を有効に防止することができる。

【0126】

以上の構成によれば、新規なユーザとの対話を通してそのユーザの名前を取得し、当該名前を、マイクロホン51やCCDカメラ50の出力に基づいて検出したそのユーザの声の音響的特徴及び顔の形態的特徴の各データと関連付けて記憶すると共に、これら記憶した各種データに基づいて、名前を取得していないさらに新規なユーザの登場を認識し、その新規なユーザの名前や声の音響的特徴及び顔の形態的特徴を上述と同様にして取得し記憶するようにしてユーザの名前を学習するようにしたことにより、学習していることをそのユーザに認識させることなく、人間が普段行うように、ユーザとの対話を通じて新規のユーザの名前やそのユーザの声の音響的特徴及び顔の形態的特徴を自然に学習することができ、かくしてエンターテインメント性を格段的に向上し得るロボットを実現できる。

【0127】

(4) 他の実施の形態

なお上述の実施の形態においては、本発明を図1のように構成された2足歩行型のロボット1に適用するようにした場合について述べたが、本発明はこれに限らず、この他種々の形態のロボット装置及びロボット装置以外のこの他種々の装置に広く適用することができる。

【0128】

また上述の実施の形態においては、学習対象が人間（ユーザ）である場合につ

いて述べたが、本発明はこれに限らず、人間以外の物体を名前学習の対象とする場合においても適用することができる。

【0129】

この場合において、上述の実施の形態においては、対象となる人物の声の音響的特徴及び顔の形態的特徴からその人物をそれぞれ認識し、これらの認識結果に基づいてその人物が新規な人物であるか否かを判別するようにした場合について述べたが、本発明はこれに限らず、これに代えて又はこれに加えて、これら以外の例えば体型やにおい等の生物学的に個体を識別可能な複数種類の各種特徴からその人物をそれぞれ認識し、これらの認識結果に基づいてその人物が新規な人であるか否かを判別するようにしても良い。また名前学習対象が人物以外の物体である場合には、色や形状、模様、大きさ等の物体を識別可能な複数種類の特徴からそれぞれその物体を認識し、これらの認識結果に基づいてその物体が新規な物体であるか否かを判別するようにしても良い。そしてこれらの場合には、それぞれ物体の異なる所定の特徴を検出すると共に、当該検出結果及び予め記憶している既知の物体の対応する特徴のデータに基づいて、当該対象とする物体を認識する複数の認識手段を設けるようにすれば良い。

【0130】

さらに上述の実施の形態においては、学習達成度を「A」～「C」の3段階とするようにした場合について述べたが、本発明はこれに限らず、2段階又は4段階以上とするようにしても良い。

【0131】

さらに上述の実施の形態においては、例えば名前学習処理手順R T 1 のステップS P 1 0～ステップS P 1 3について上述した追加学習時に、対話制御部6 1 が単に対話を引き伸ばすだけである場合について述べたが、本発明はこれに限らず、ユーザとの対話を引き伸ばす際に、対話制御部6 3が、そのユーザに対する学習が不十分な認識手段（話者認識部6 1 及び顔認識部6 2）が学習し易いような対話をユーザとの間で生成するように処理するようにしても良く、これにより追加学習をより効率良く行うことができる。

【0132】

実際上、例えば実施の形態の場合であれば、追加学習を行うのが話者認識部61である場合には、なるべくユーザに発話させるようにな対話を生成しながら対話を引き伸ばし、追加学習を行うのが顔認識部62である場合には、なるべくユーザのいろいろな方向からの顔画像を取得できるように、「右を向いてくれる？」など、ユーザの顔を動かさせるような対話を生成しながら対話を引き伸ばすようすれば良い。

【0133】

さらに上述の実施の形態においては、連想記憶部65が、記憶した各学習達成度を例えれば数日ごとに時間減衰させるようにした場合について述べたが、本発明はこれに限らず、時間減衰させる時間間隔は数日単位以外であっても良く、またかかる時間減衰を対話制御部63が管理して行わせるようにしても良い。

【0134】

さらに上述の実施の形態においては、ユーザとの対話を通して対象とする物体の名前をユーザから取得する対話手段と、当該対話手段が取得した対象とする物体の名前、当該対象とする物体に対する各認識手段の認識結果、及び記憶手段が記憶する関連付け情報に基づいて、対象とする物体が新規であると判断したときには、必要な認識手段に当該対象とする物体の対応する特徴を学習させると共に、当該対象とする物体についての関連付け情報を記憶手段に新たに記憶させる制御手段とを同じ1つの機能モジュールである対話制御部63により構成するようにした場合について述べたが、本発明はこれに限らず、これらを別のモジュールとして構成するようにしても良い。

【0135】

さらに上述の実施の形態においては、既知のユーザの名前及び当該既知のユーザに対する各認識部（音声認識部60、話者認識部61及び顔認識部62）の認識結果を関連付けた関連付け情報を記憶する記憶手段を、内部メモリ40Aとソフトウェアとから構成するようにした場合について述べたが、本発明はこれに限らず、例えればかかる記憶手段における関連付け情報を記憶する機能部分を、内部メモリ40A以外の例ええばコンパクトディスク等の読み書き自在の他の記憶機能を有する手段に置き換えるようにしても良い。

【0136】

【発明の効果】

上述のように本願発明によれば、学習装置において、ユーザとの対話を通して対象とする物体の名前をユーザから取得する対話手段と、それぞれ対象とする物体の異なる所定の特徴を検出すると共に、当該検出結果及び予め記憶している既知の物体の対応する特徴の学習結果に基づいて、当該対象とする物体を認識する複数の認識手段と、既知の物体の名前及び当該既知の物体に対する各認識手段の認識結果を関連付けた関連付け情報を記憶する記憶手段と、対話手段が取得した対象とする物体の名前、当該対象とする物体に対する各認識手段の認識結果、及び記憶手段が記憶する関連付け情報に基づいて、対象とする物体が新規であると判断したときには、必要な認識手段に当該対象とする物体の対応する特徴を学習させると共に、当該対象とする物体についての関連付け情報を記憶手段に新たに記憶させる制御手段とを設けるようにしたことにより、ユーザにそれと認識されることなく、対象とする物体の名前を通常のユーザとの対話を通して学習することができ、かくしてエンターテインメント性を向上させ得る学習装置を実現できる。

【0137】

また本発明においては、学習方法において、ユーザとの対話を通して当該対象とする物体の名前をユーザから取得すると共に、対象とする物体の複数の異なる所定の特徴についての検出結果及び予め記憶している既知の物体の各特徴についての学習結果に基づいて、当該対象とする物体を認識する第1のステップと、取得した対象とする物体の名前と、当該対象とする物体の各特徴にそれぞれ基づく認識結果と、予め記憶している既知の物体の名前、当該既知の物体の各特徴に対する認識結果を関連付けた関連付け情報を基づいて、対象とする物体が新規であると判断したときには、当該対象とする物体の必要な特徴を学習すると共に、当該対象とする物体についての関連付け情報を新たに記憶する第2のステップとを設けるようにしたことにより、ユーザにそれと認識されることなく、対象とする物体の名前を通常のユーザとの対話を通して学習することができ、かくしてエンターテインメント性を向上させ得る学習方法を実現できる。

【0138】

さらに本発明においては、ロボット装置において、ユーザとの対話を通して対象とする物体の名前をユーザから取得する対話手段と、それぞれ対象とする物体の異なる所定の特徴を検出すると共に、当該検出結果及び予め記憶している既知の物体の対応する特徴の学習結果に基づいて、当該対象とする物体を認識する複数の認識手段と、既知の物体の名前及び当該既知の物体に対する各認識手段の認識結果を関連付けた関連付け情報を記憶する記憶手段と、対話手段が取得した対象とする物体の名前、当該対象とする物体に対する各認識手段の認識結果、及び記憶手段が記憶する関連付け情報に基づいて、対象とする物体が新規であると判断したときには、必要な認識手段に当該対象とする物体の対応する特徴を学習させると共に、当該対象とする物体についての関連付け情報を記憶手段に新たに記憶させる制御手段とを設けるようにしたことにより、ユーザにそれと認識されることなく、対象とする物体の名前を学習することができ、かくしてエンターテイメント性を向上させ得るロボット装置を実現できる。

【図面の簡単な説明】**【図1】**

本実施の形態によるロボットの外観構成の説明に供する斜視図である。

【図2】

本実施の形態によるロボットの外観構成の説明に供する斜視図である。

【図3】

本実施の形態によるロボットの外観構成の説明に供する概念図である。

【図4】

本実施の形態によるロボットの内部構成の説明に供するブロック図である。

【図5】

本実施の形態によるロボットの内部構成の説明に供するブロック図である。

【図6】

名前学習機能に関するメイン制御部の機能の説明に供するブロック図である。

【図7】

連想記憶部における各種情報の関連付けの説明に供する概念図である。

【図8】

名前学習処理手順を示すフローチャートである。

【図9】

ロボットとユーザとの対話例を示す略線図である。

【図10】

ロボットとユーザとの対話例を示す略線図である。

【図11】

エラー処理手順を示すフローチャートである。

【図12】

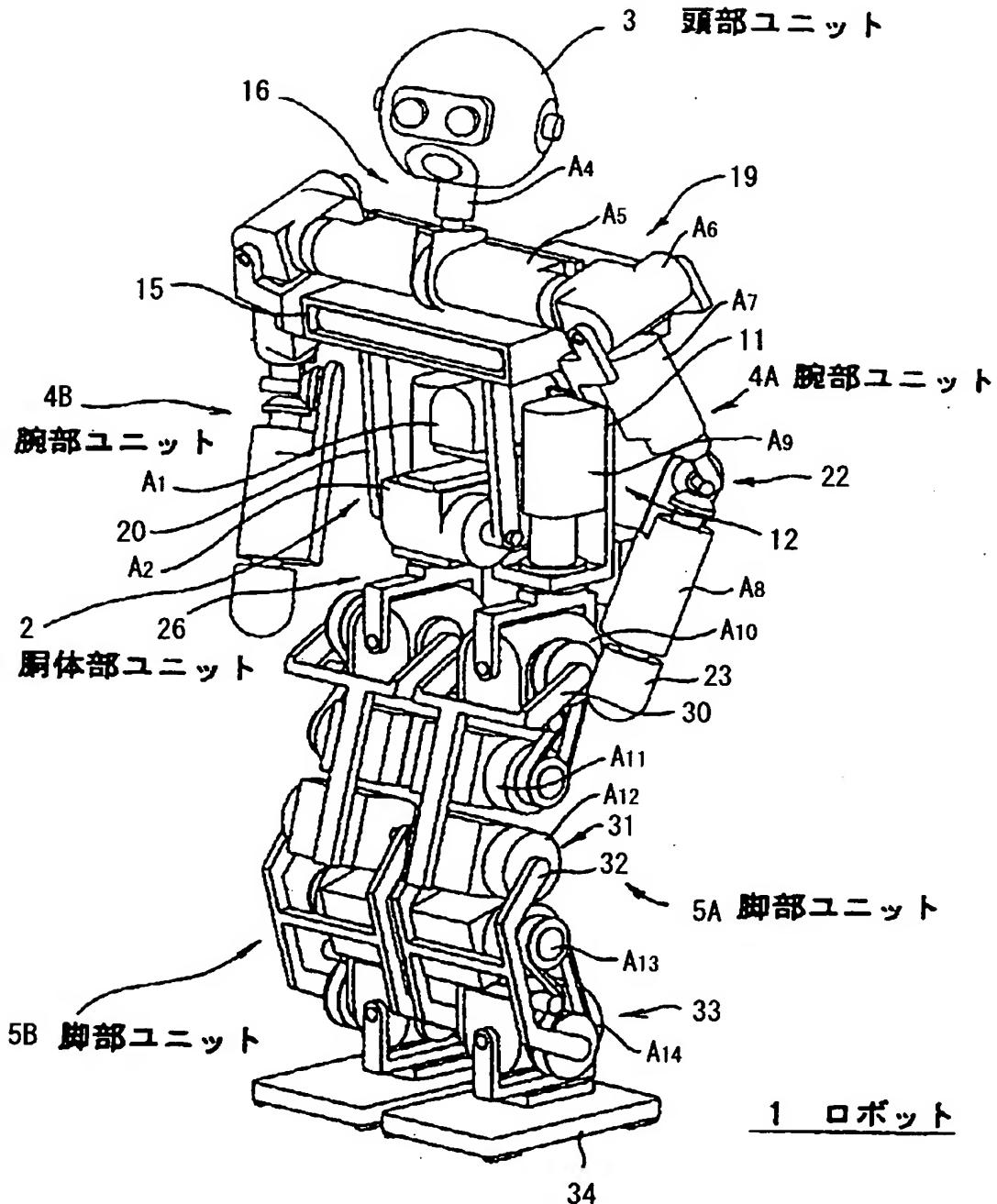
ロボットとユーザとの対話例を示す略線図である。

【符号の説明】

1……ロボット、 4 0……メイン制御部、 5 0……CCDカメラ、 5 1……マイクロホン、 5 2……スピーカ、 6 0……音声認識部、 6 1……話者認識部、 6 2……顔認識部、 6 3……対話制御部、 6 4……音声合成部、 6 5……連想記憶部、 S 1 A……画像信号、 S 1 B、 S 3……音声信号、 D 1、 D 2……文字列データ、 R T 1……名前学習処理手順、 R T 2……エラー処理手順。

【書類名】 図面

【図1】



【図2】

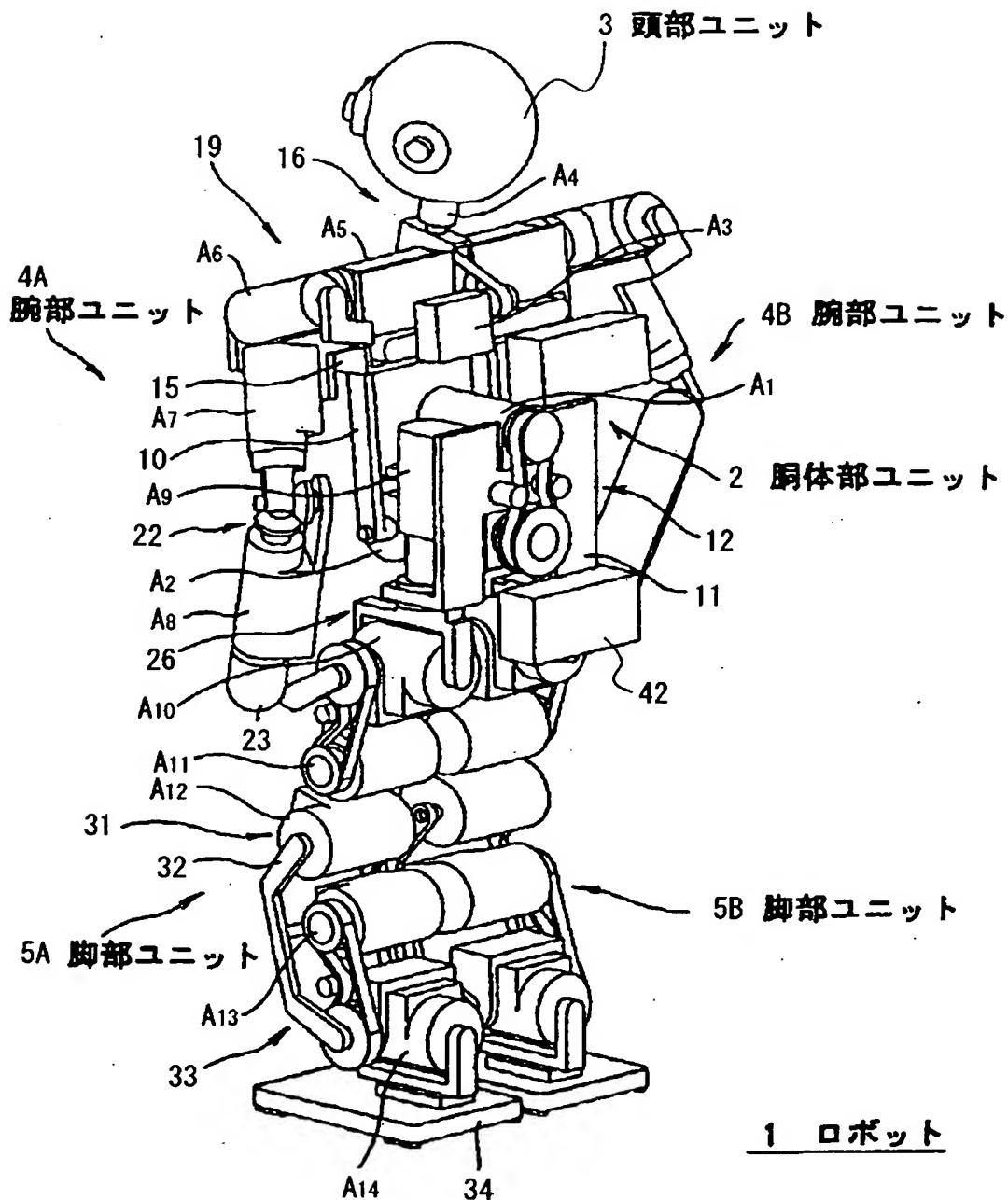


図2 本実施の形態によるロボットの外観構成（2）

【図3】

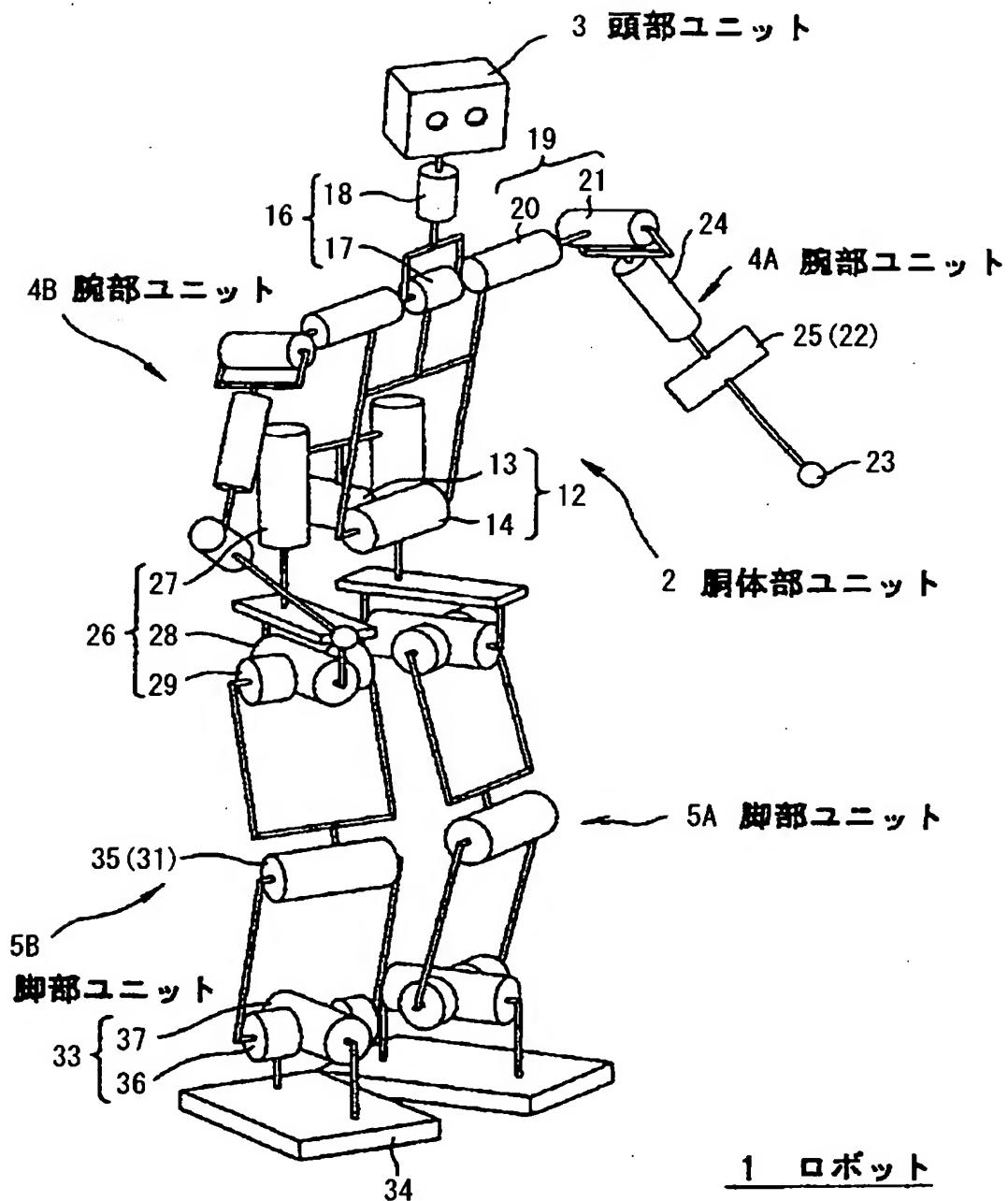


図3 本実施の形態によるロボットの外観構成（3）

【図4】

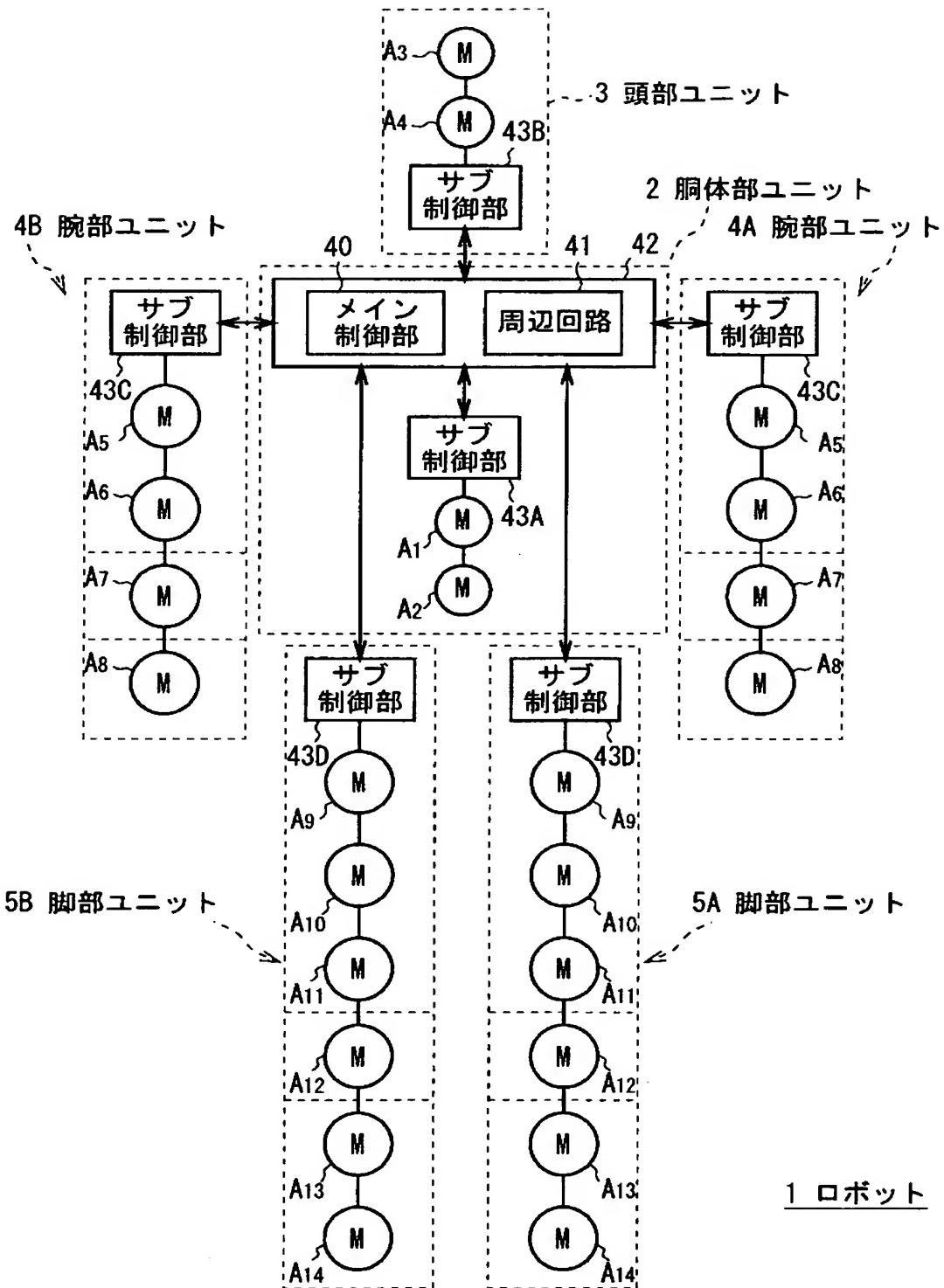


図4 ロボットの内部構成（1）

【図5】

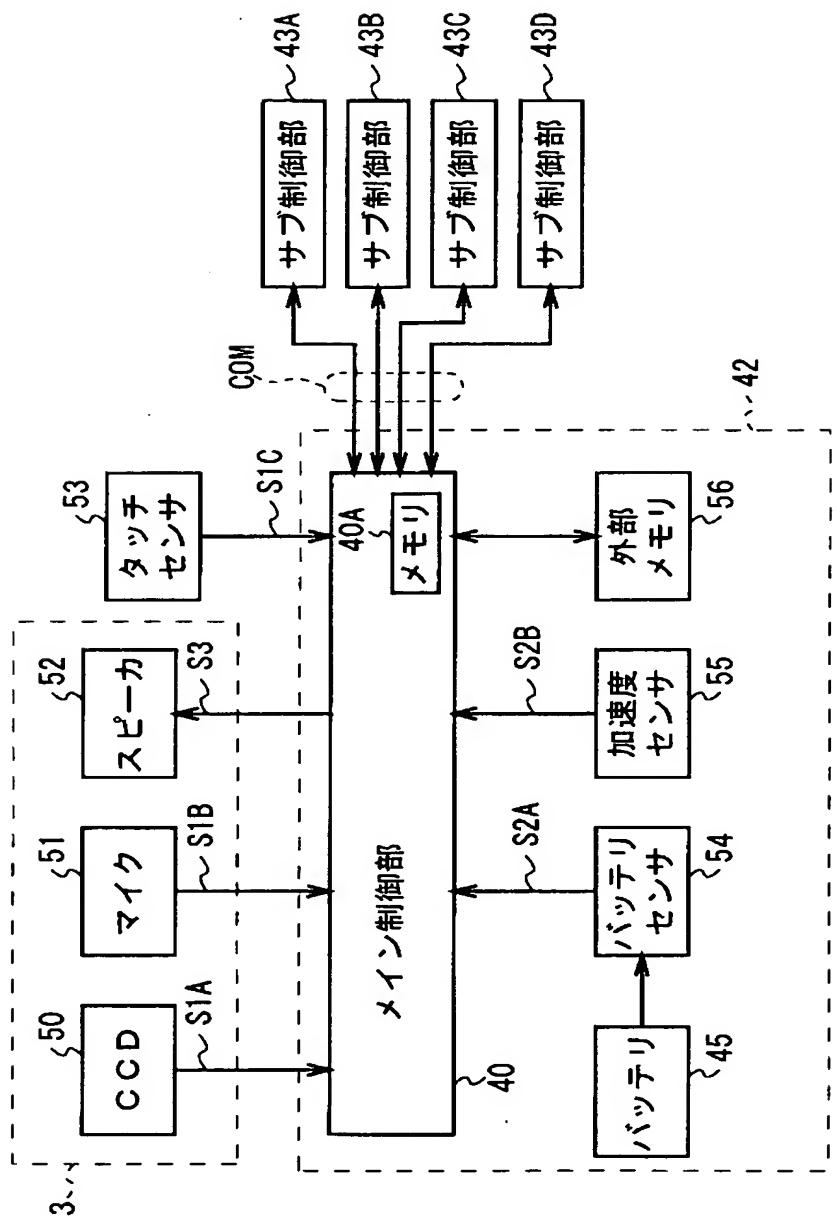


図5 ロボットの内部構成(2)

【図6】

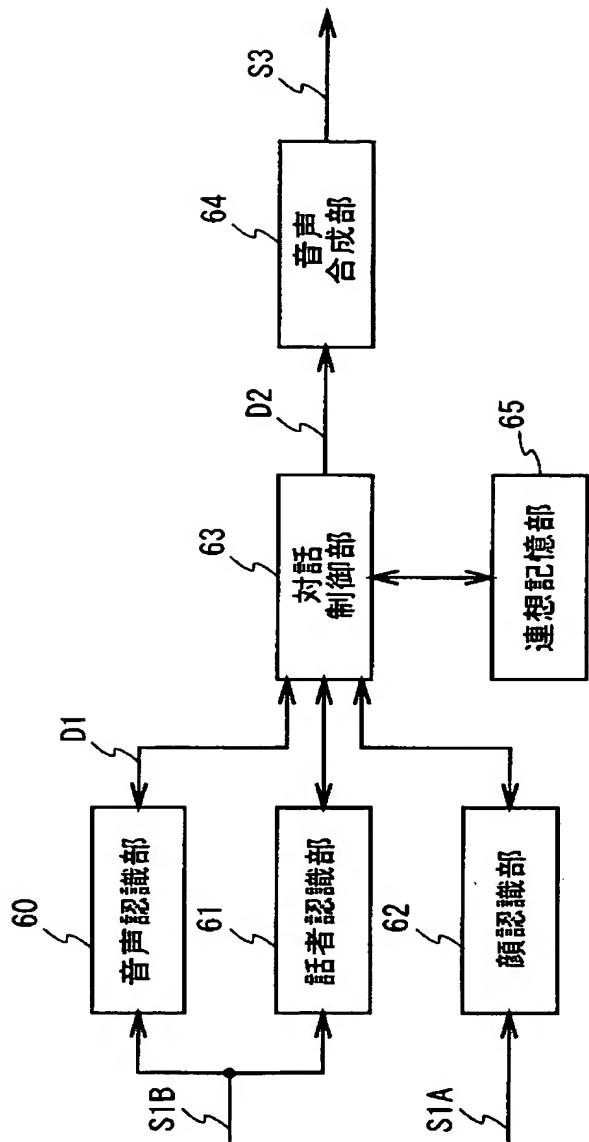


図6 メイン制御部の処理

【図7】

F I D		S I D		名前
1	A	2	B	フジタ
2	C	5	B	ヨシダ

図7 連想記憶部における各情報の関連付け

【図8】

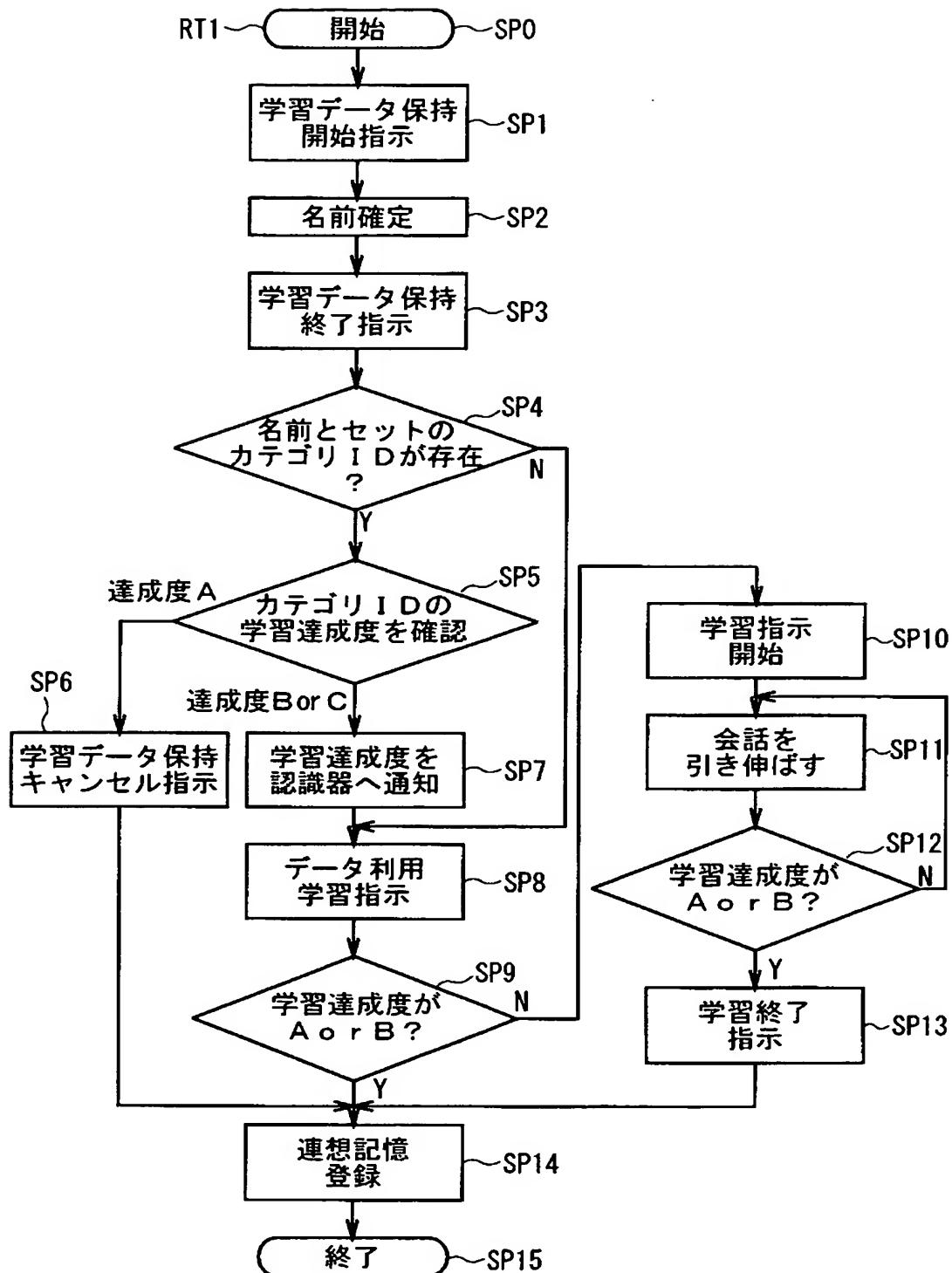


図8 名前学習処理順

【図9】

ロボット：名前を教えてください。
ユーザ：私の名前はヤマモトです
ロボット：ヤマモトさんっていうんですね
ユーザ：そうだよ
ロボット：こんにちは！ヤマモトさん！

図9 ユーザとの対話例（1）

【図10】

ロボット：僕とお友達になってくれる？
ユーザ：いいよ。
ロボット：ありがとう！それじゃ、ヤマモトさんのこと聞いていい？
ユーザ：いいよ
ロボット：ヤマモトさんの好きなたべものって何？

図10 ユーザとの対話例（2）

【図11】

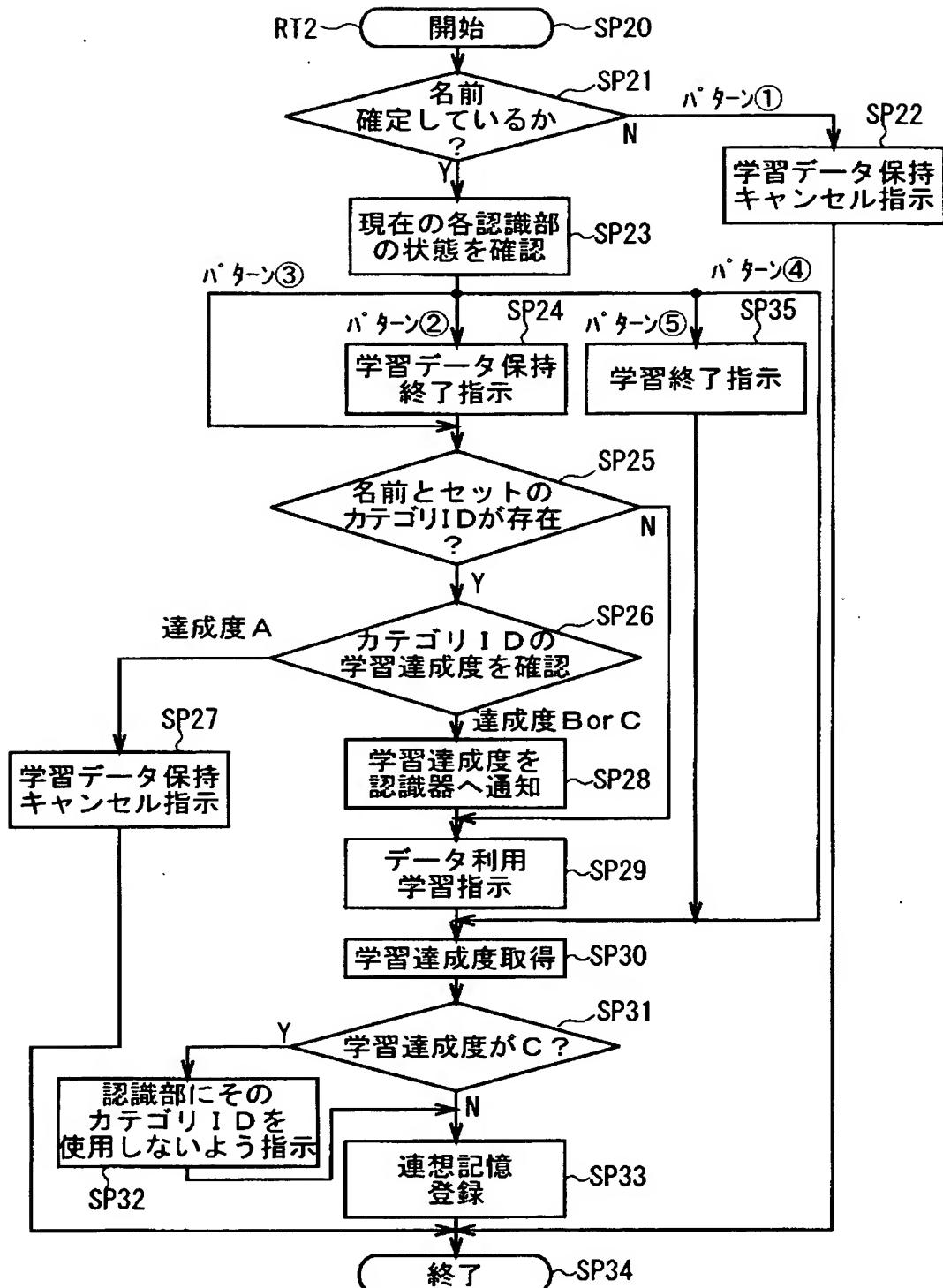


図11 エラー処理手順

【図12】

ロボット：名前を教えてください。
ユーザ：私の名前はヤマモトです
ロボット：ヤマモトさんっていうんですね
ユーザ：そうだよ
ロボット：ヤマモトさん！あ、前にお会いしたことありますね！

図12 ユーザとの対話例（3）

【書類名】 要約書

【要約】

【課題】

エンターテインメント性を向上させ得る学習装置及び学習方法並びにロボット装置を提案する。

【解決手段】

ユーザとの対話を通して当該対象とする物体の名前をユーザから取得すると共に、対象とする物体の複数の異なる所定の特徴についての検出結果及び予め記憶している既知の物体の各特徴についての学習結果に基づいて、当該対象とする物体を認識し、取得した対象とする物体の名前と、当該対象とする物体の各特徴にそれぞれ基づく認識結果と、予め記憶している既知の物体の名前、当該既知の物体の各特徴に対する認識結果を関連付けた関連付け情報に基づいて、対象とする物体が新規であると判断したときには、当該対象とする物体の必要な特徴を学習すると共に、当該対象とする物体についての関連付け情報を新たに記憶するようにした。

【選択図】 図8

特願 2003-041833

出願人履歴情報

識別番号 [000002185]

1. 変更年月日 1990年 8月30日
[変更理由] 新規登録
住 所 東京都品川区北品川6丁目7番35号
氏 名 ソニー株式会社